

DATA ASSIMILATION FOR PM25 RIO MAPS

08/10/2021

Jorge Sousa, Stijn Janssen, Stijn Vranckx

☐ Air quality model: RIO

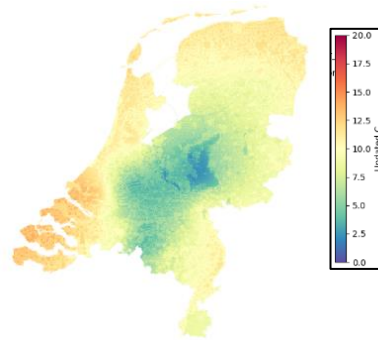
- Geospatial detrended ordinary kriging interpolation
- Uncertainty in the concentration values is available
- Hourly maps output

☐ Low-cost sensor network:

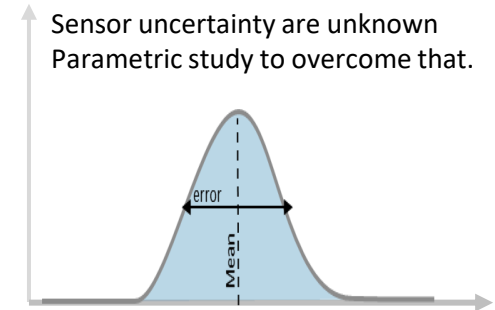
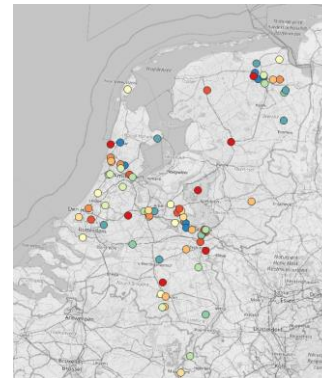
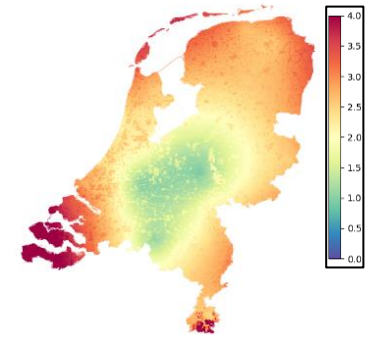
1. Calibration
2. Data Fusion

- Sub-set: 80 sensors: <https://sensors.rivm.nl/>
- Uncertainties are unknown at this moment
 1. Sensitivity to uncertainty of sensors
 2. Sensitivity to sensor density

Concentration map PM25 [$\mu\text{g}/\text{m}^3$]

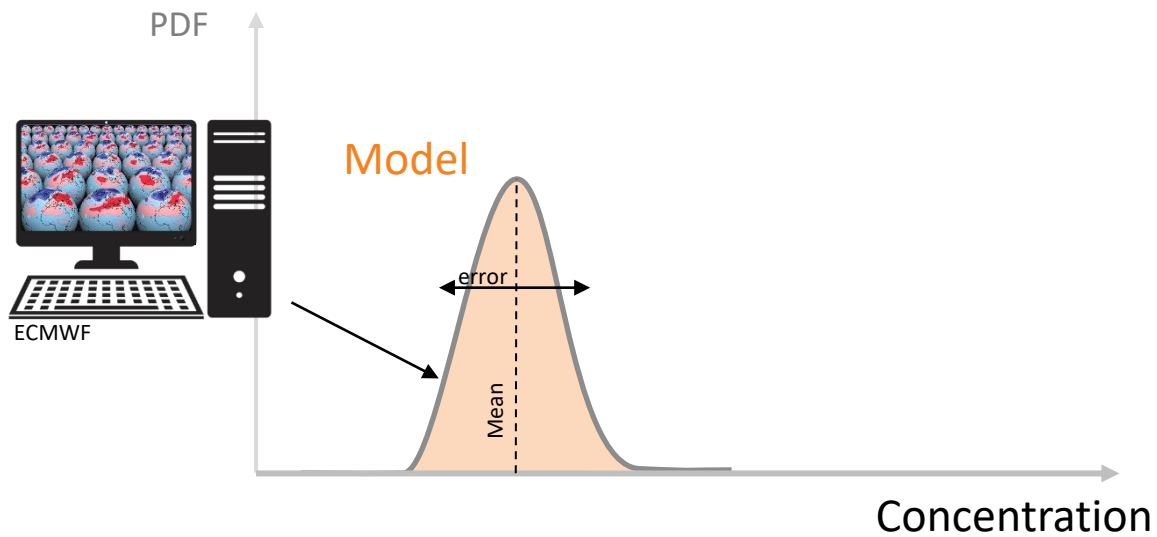


Uncertainty map PM25 [$\mu\text{g}/\text{m}^3$]

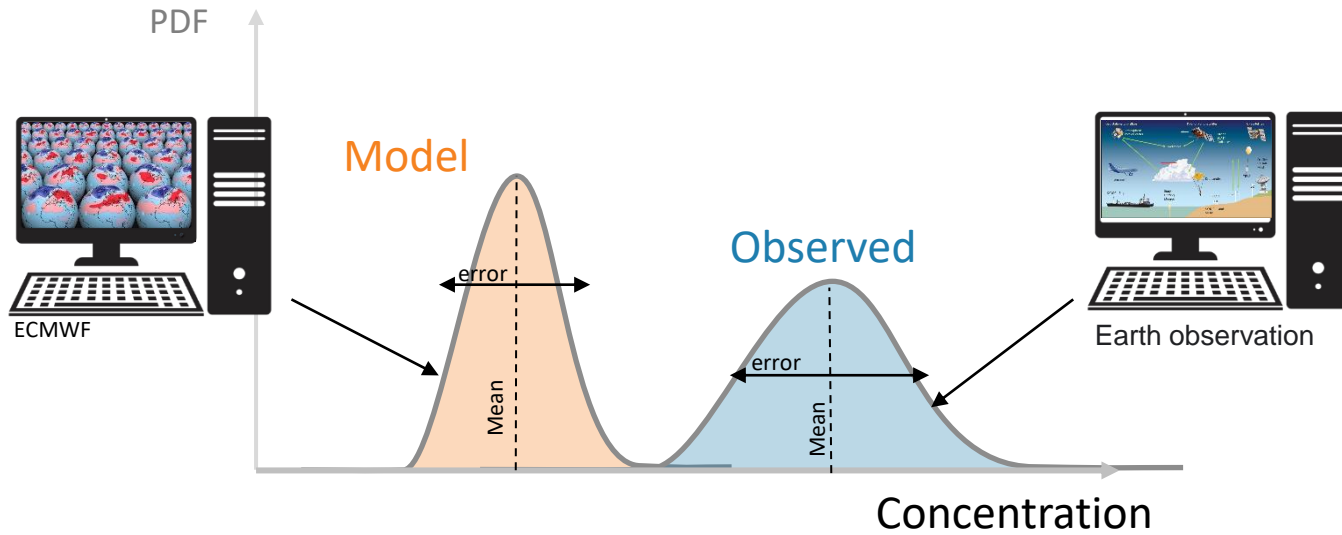


Objective: Identity the possibilities to update maps with highly uncertain sensors

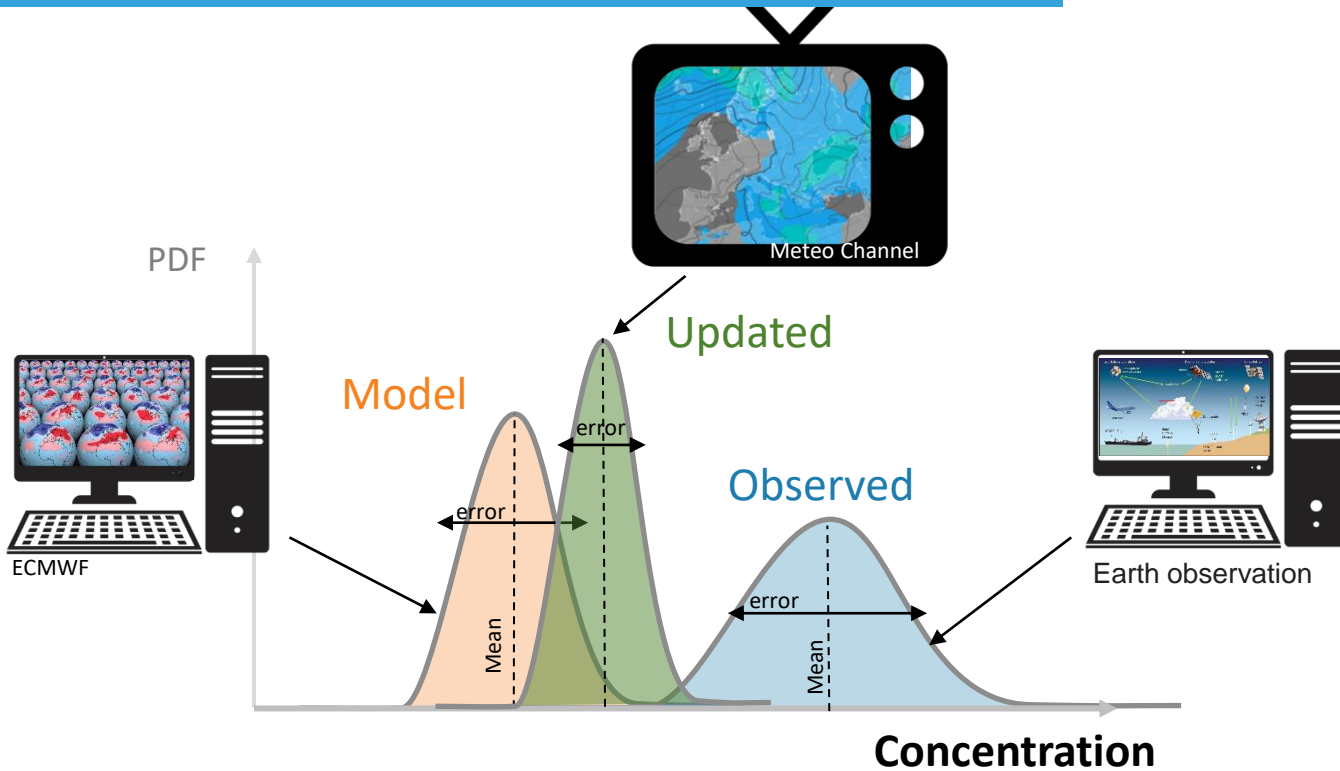
DATA ASSIMILATION: EXAMPLE FROM WEATHER FORECAST



DATA ASSIMILATION: EXAMPLE FROM WEATHER FORECAST



DATA ASSIMILATION: EXAMPLE FROM WEATHER FORECAST

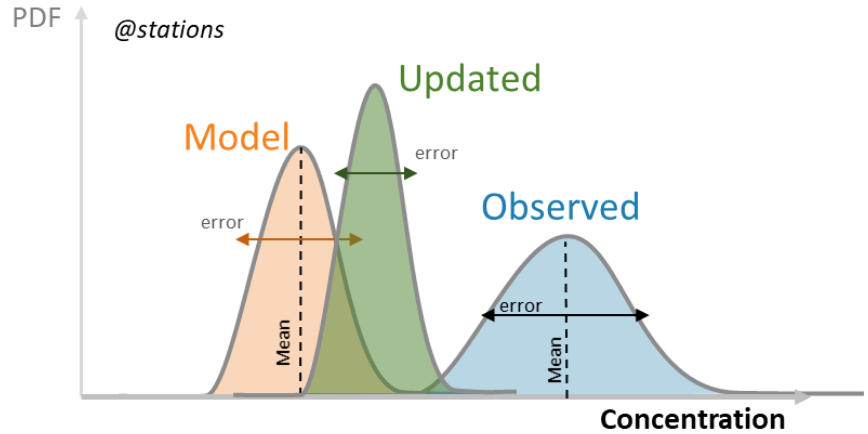


- Updated state merges model and observation results by minimizing its variance (uncertainty)

DATA ASSIMILATION:

Bayesian approach for map fusion:

- Considers both model and measurement uncertainty
- Updates concentration and uncertainty values of fused map
- BLUE (Best Linear Unbiased Estimator) – Kalman Filter



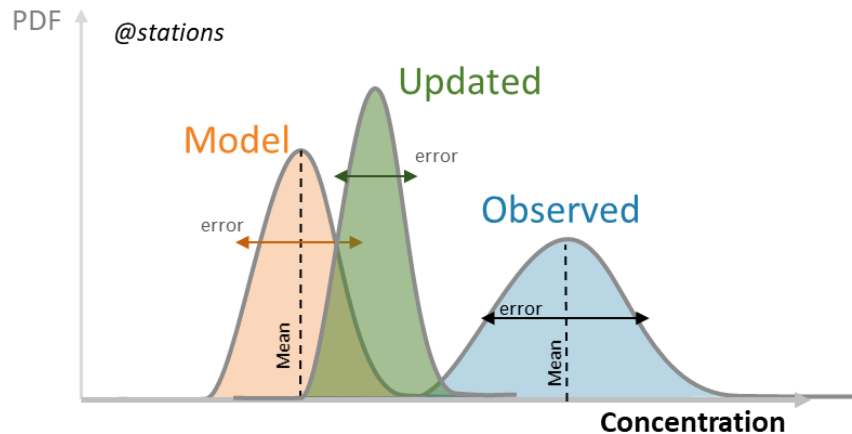
DATA ASSIMILATION:

Bayesian approach for map fusion:

- Considers both model and measurement uncertainty
- Updates concentration and uncertainty values of fused map
- BLUE (Best Linear Unbiased Estimator) – Kalman Filter

1) Concentration update

$$\overset{\text{Updated map}}{\psi^a} = \overset{\text{Initial map}}{\psi^f} + K \left(\underset{\text{Obs.}}{d} - \overset{\text{Model @stations}}{M\psi^f} \right)$$



DATA ASSIMILATION:

Bayesian approach for map fusion:

- Considers both model and measurement uncertainty
- Updates concentration and uncertainty values of fused map
- BLUE (Best Linear Unbiased Estimator) – Kalman Filter

1) Concentration update

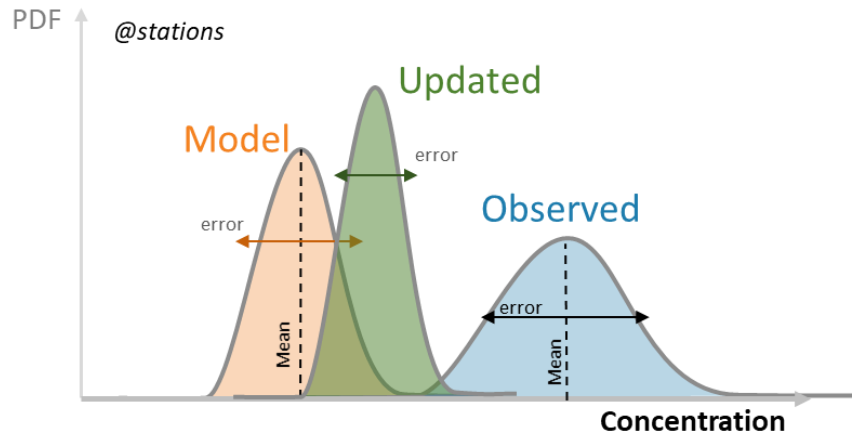
$$\psi^a = \psi^f + K(d - M\psi^f)$$

Updated map Initial map Obs. Model @stations

2) Error update

$$C_{\psi\psi}^a = (I - KM)C_{\psi\psi}^f$$

Updated Covariance matrix Initial Covariance matrix



DATA ASSIMILATION:

Bayesian approach for map fusion:

- Considers both model and measurement uncertainty
- Updates concentration and uncertainty values of fused map
- BLUE (Best Linear Unbiased Estimator) – Kalman Filter

1) Concentration update

$$\psi^a = \psi^f + K(d - M\psi^f)$$

Updated map Initial map Obs. Model @stations

2) Error update

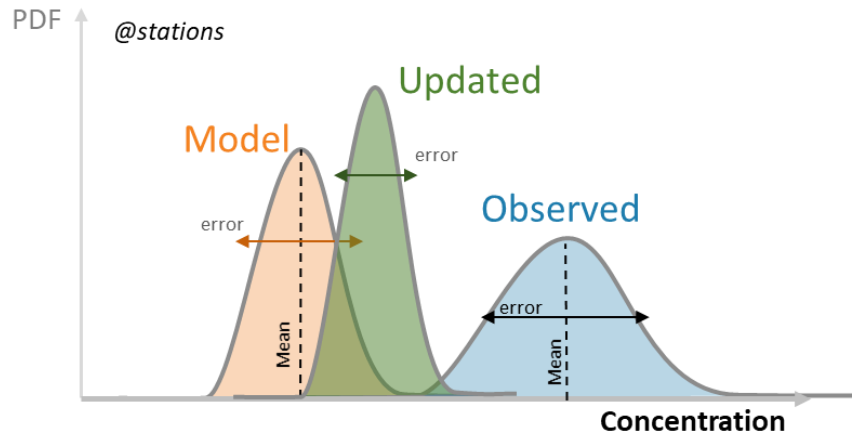
$$C_{\psi\psi}^a = (I - KM)C_{\psi\psi}^f$$

Updated Covariance matrix Initial Covariance matrix

3) Kalman Gain

$$K = C_{\psi\psi}^f M^T (M C_{\psi\psi}^f M^T + C_{\varepsilon\varepsilon})^{-1}$$

Kalman Gain Measurement covariance



DATA ASSIMILATION:

Bayesian approach for map fusion:

- Considers both model and measurement uncertainty
- Updates concentration and uncertainty values of fused map
- BLUE (Best Linear Unbiased Estimator) – Kalman Filter

1) Concentration update

$$\psi^a = \psi^f + K(d - M\psi^f)$$

Updated map Initial map Obs. Model @stations

2) Error update

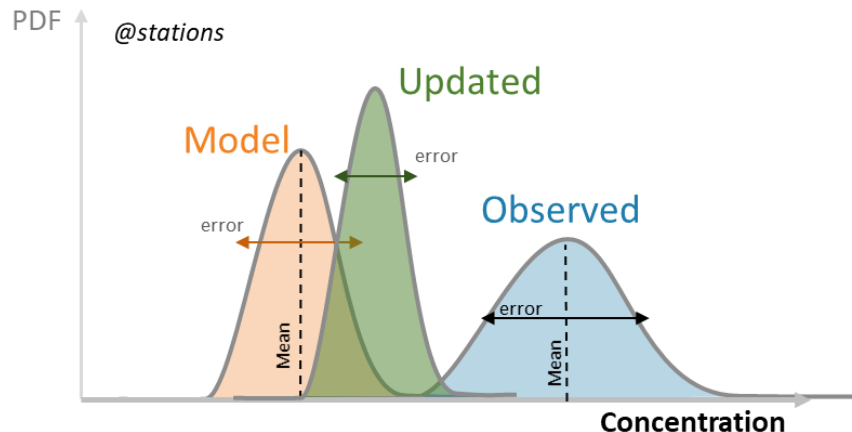
$$C_{\psi\psi}^a = (I - KM)C_{\psi\psi}^f$$

Updated Covariance matrix Initial Covariance matrix

3) Kalman Gain

$$K = C_{\psi\psi}^f M^T (M C_{\psi\psi}^f M^T + C_{\varepsilon\varepsilon})^{-1}$$

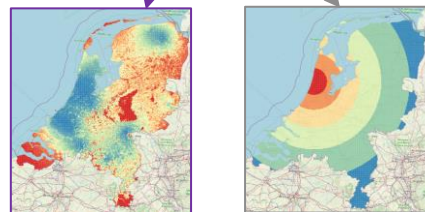
Kalman Gain Measurement covariance



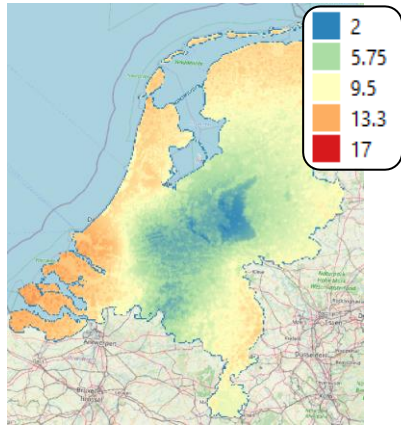
4) State covariance matrix

$$C_{\psi\psi}^f = E a e^{-L\psi\psi/\tau}$$

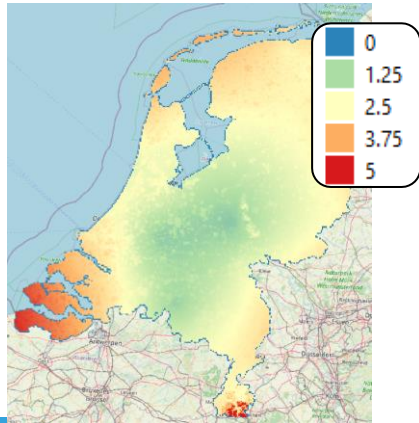
RIO error



Initial

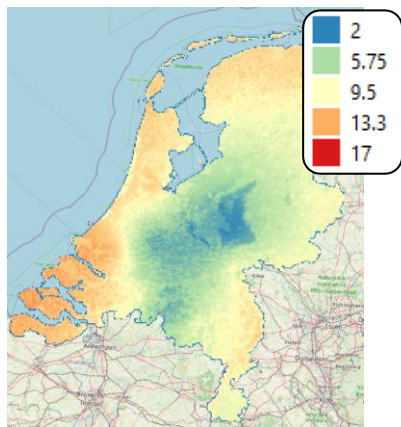


Concentration
PM25
[$\mu\text{g}/\text{m}^3$]

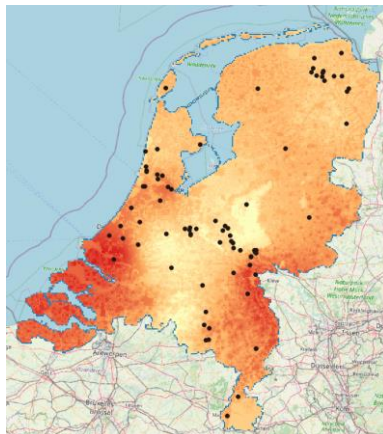


Error
PM25
[$\mu\text{g}/\text{m}^3$]

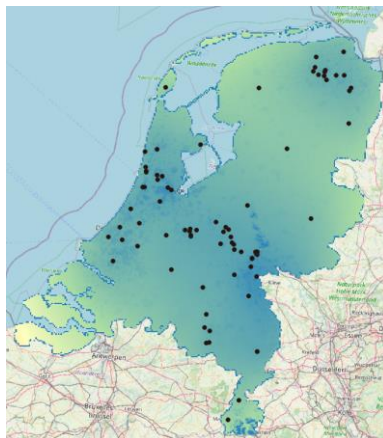
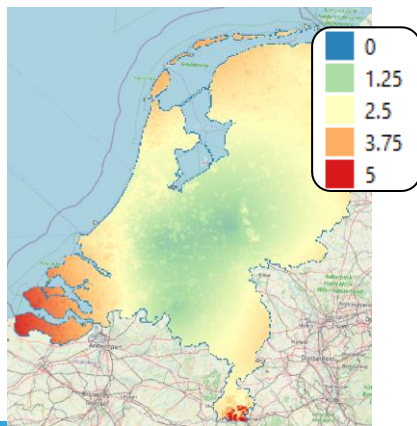
Initial



Updated: 50% obs. uncertainty

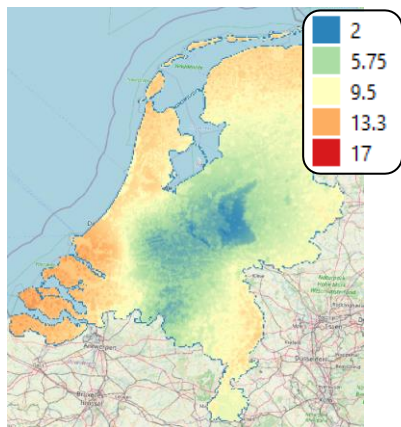


Concentration
PM25
[ug/m³]



Error
PM25
[ug/m³]

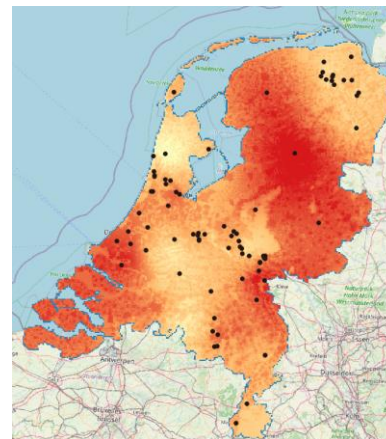
Initial



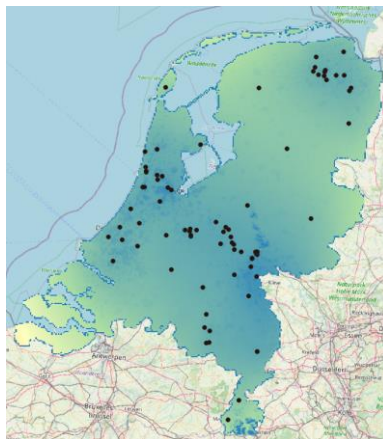
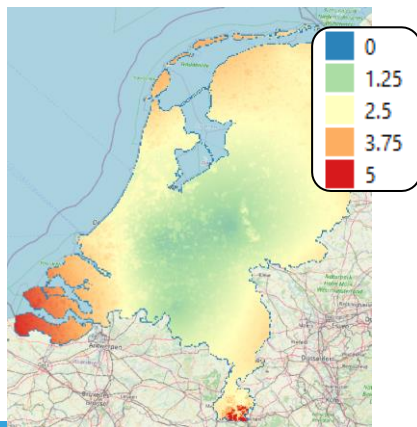
Updated: 50% obs. uncertainty



Updated: 10% obs. uncertainty



Error
PM25
[ug/m³]

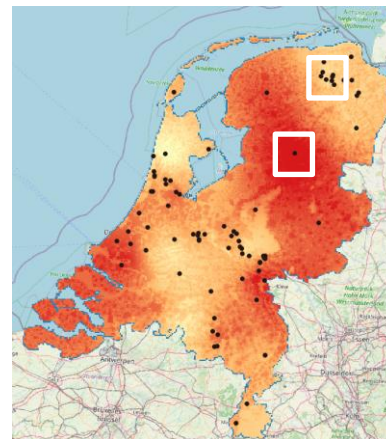
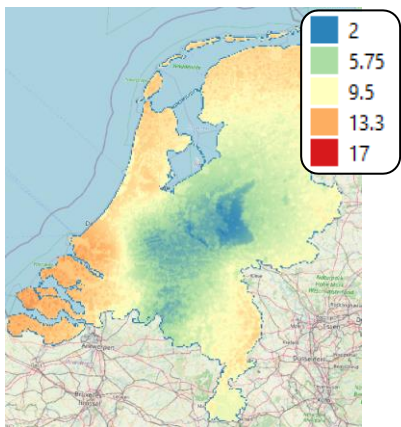


Initial

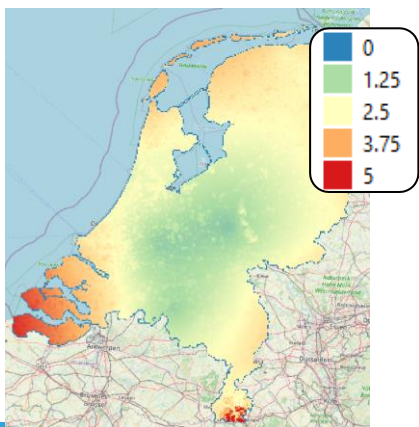
Updated: 50% obs. uncertainty

Updated: 10% obs. uncertainty

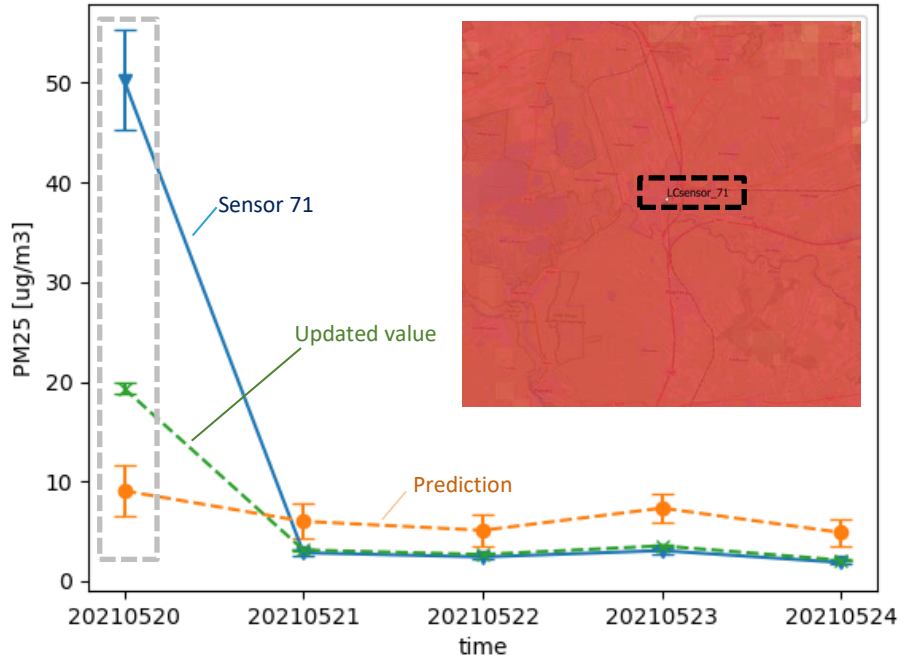
Concentration
PM25
[ug/m³]



Error
PM25
[ug/m³]

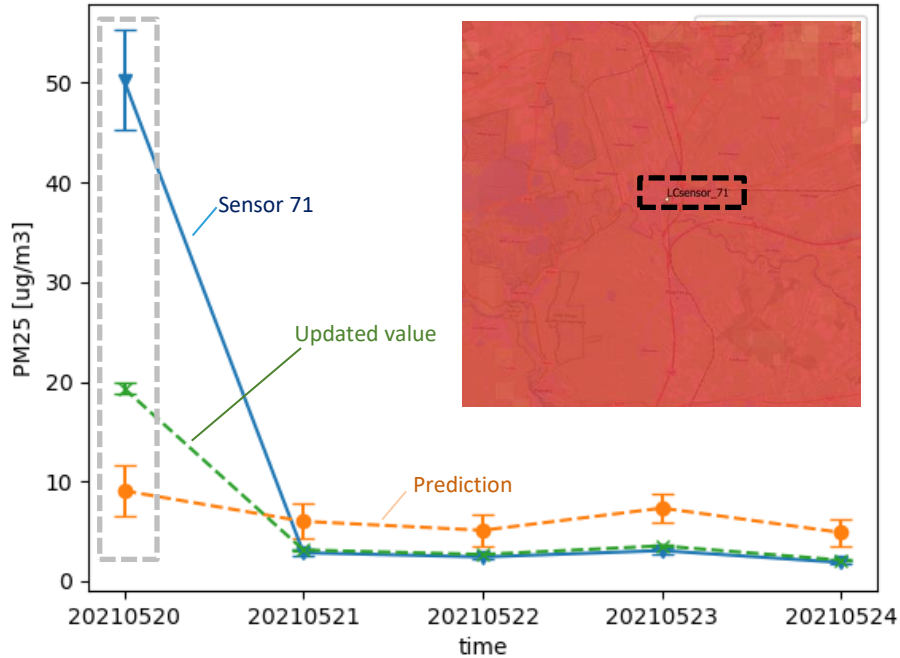


Sensor 71 - Isolated



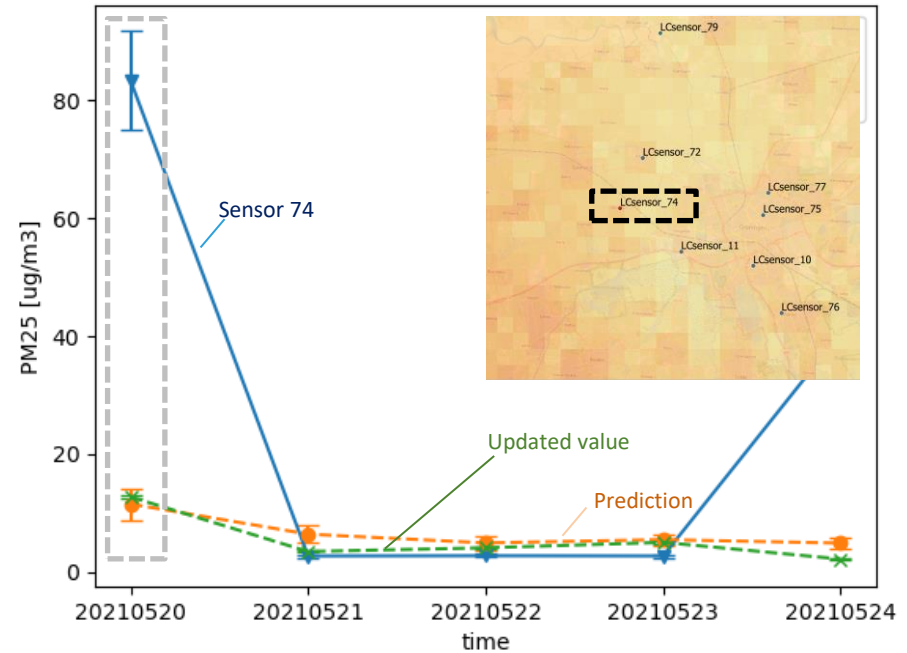
- Signification weight given to sensor 71
- Outliers more critical when isolated

Sensor 71 - Isolated



- Signification weight given to sensor 71
- Outliers more critical when isolated

Sensor 74 – not isolated



- Few weight given to sensor 74
- Statistical less significant because of the surrounding sensors

Take away messages:

Bayesian approach

- Considers both model and sensor uncertainty
- Updates mean values and uncertainty

Impact of sensor uncertainty:

- Model and observation uncertainty are compared and more weight is given to the lowest of the 2
- The knowledge of the uncertainty for each sensor and for each hour is typically necessary

Sensor density:

- The data fusion methodologies can provide less statistical significance to local outliers.
- Problematic when sensors are isolated.